WPS8433

Policy Research Working Paper 8433

Gender Gap in Earnings in Vietnam

Why Do Vietnamese Women Work in Lower Paid Occupations?

Iffat Chowdhury
Hillary C. Johnson
Aneesh Mannava
Elizaveta Perova

WORLD BANK GROUP
East Asia and the Pacific Region
Office of the Chief Economist
May 2018

Policy Research Working Paper 8433

Abstract

Differences in earnings between male and female workers persist in developed and developing countries despite a narrowing of gender gaps in educational attainment over the past half-century. This paper examines the gender wage gap in Vietnam and shows that a nontrivial part of the gap is associated with occupational sorting. The paper considers three explanations for why occupational sorting emerges. First, it explores whether women sort into occupations with better nonmonetary characteristics, such as paid leave and shorter hours. The data from Labor Force Surveys support this hypothesis. Second, it checks if occupational sorting among the adult labor force is driven by social norms about gender roles learned and internalized at an early age. To do so, the paper checks for evidence of sorting in the aspirations of 12-year-old children. Specifically, the analysis simulates what the gender wage gap would be if boys and girls pursued the occupations they aspired to at age 12, and the distribution of salaries remained unchanged. The paper does not find support for the hypothesis that gender norms drive occupational sorting by inducing aspirational sorting at an early age. Finally, for individuals with higher education, the paper checks if occupational sorting occurs during the school-to-work transition, when women face higher barriers in finding a job in their field of study. The analysis does not find evidence to support this last hypothesis. Overall, the findings suggest that in Vietnam gender-specific preferences for nonmonetary job characteristics play a key role in the emergence of occupational sorting.

This paper is a product of the Office of the Chief Economist, East Asia and the Pacific Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/research. The authors may be contacted at eperova@worldbank.org.

Gender Gap in Earnings in Vietnam: Why Do Vietnamese Women Work in Lower Paid Occupations?

EAST ASIA AND PACIFIC GENDER INNOVATION LAB

Iffat Chowdhury, Hillary C. Johnson, Aneesh Mannava and Elizaveta Perova

JEL Codes: J71, J33, J31, J16, J13.

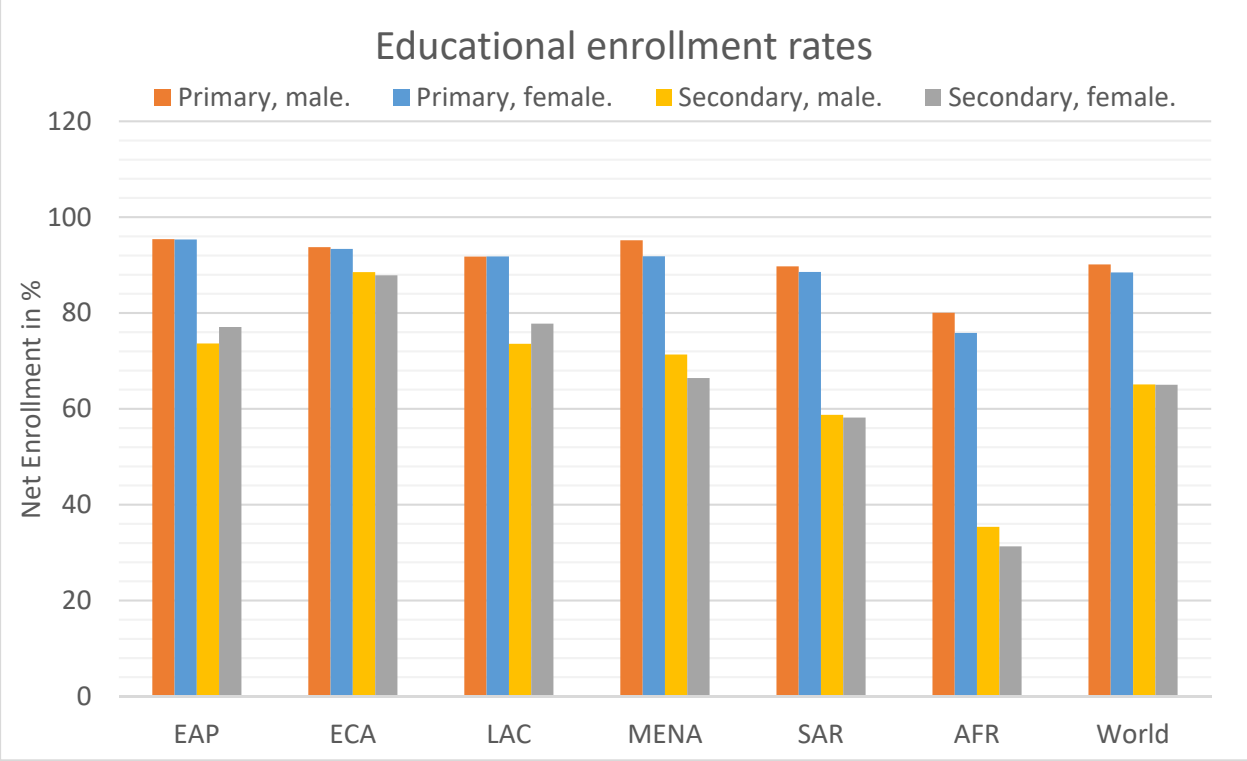
Keywords: Gender wage gap, Gender streaming, Occupational sorting, Vietnam

1. Introduction

A gap in earnings between women and men in the labor force is a common empirical feature of the labor market in countries around the world. The gender wage gap has been documented by economists for over half a century. While its magnitude has diminished in that time, it has seldom disappeared. To date, every country has some degree of wage inequality. Even in Iceland, ranked first in the World Economic Forum's Global Gender Gap Index 2016, women's earnings are on average approximately 79% that of men's. In the country with the lowest ranking, Yemen, that figure is 63%.

One factor frequently cited to explain this gap was the difference in levels of education between male and female workers. Indeed, under the Millennium Development Goal on promoting gender equality and empowering women one target was to "(e)liminate gender disparity in primary and secondary education, preferably by 2005, and in all levels of education by no later than 2015." Over the long term, perhaps in part as a result of such directed efforts, the global trend has been a narrowing of the education gap between men and women, and even a reversal in some countries including Vietnam. Figure 1 shows that male and female enrollment rates in primary and secondary education are similar, with female enrollment exceeding male enrollment at the secondary level in East Asia and Latin America.

Figure 1: Gender gaps in enrollment around the world

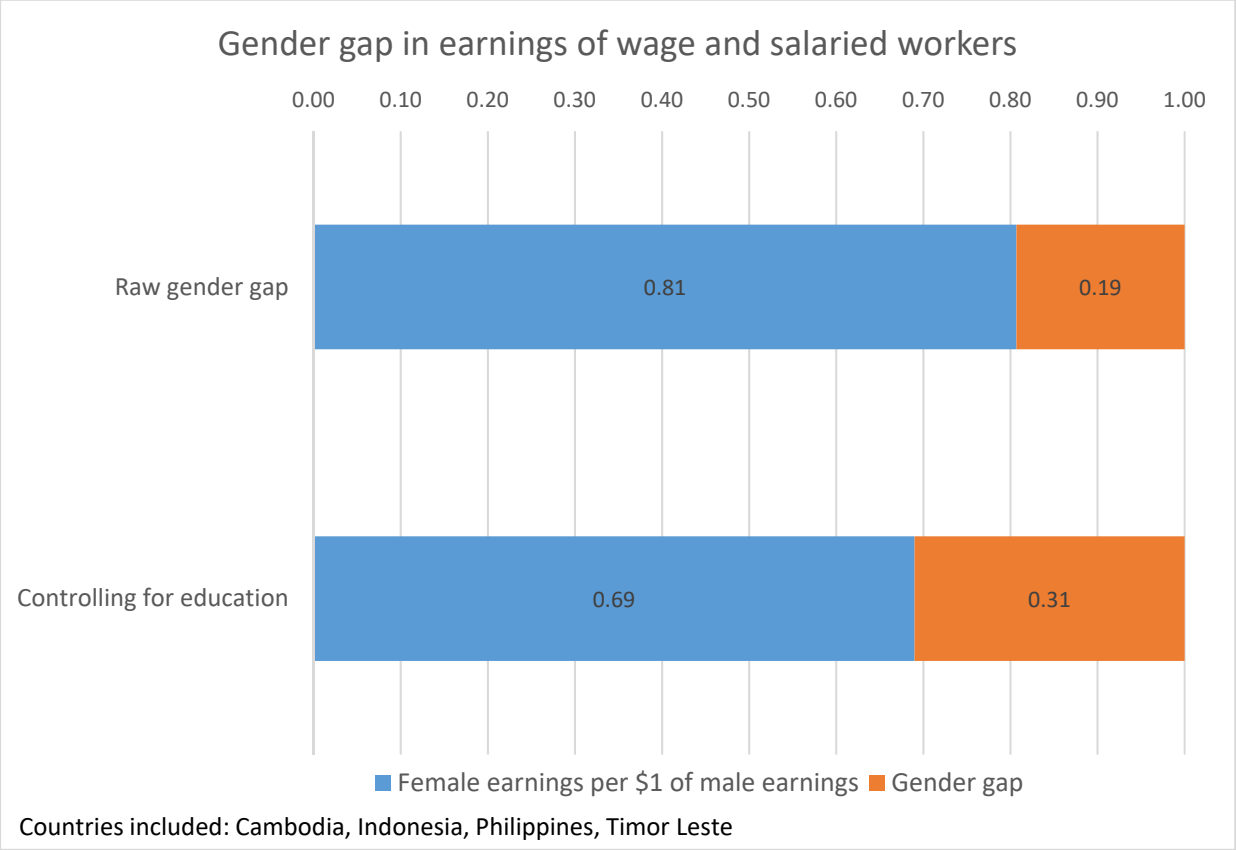


Source: World Development Indicators (WDI) 2014. The latest available date was used for every region.

The persistence of the gender wage gap stands at odds with the narrowing of the education gap. Figure 2 makes this point more strongly - for a set of countries in East Asia, not including Vietnam, which we explore in this study, controlling for levels of education actually increases the magnitude of the gender

wage gap, i.e. for women and men with the same level of education, the gender wage gap is larger than the average gender wage gap.

Figure 2: Gender wage gap widens when we control for education



Source: World Bank staff calculations using EAPPOV harmonized database

If differences in education levels do not play a role in explaining the gender wage gap in East Asian countries today, what factors are responsible for observed differences in wages between men and women? One factor that has received increasing attention in the literature is occupational sorting. Peterson and Morgan (1995) find that wage differences between men and women are very small within an establishment for the same occupation. In their study, sorting into establishments and, in particular, occupations explains a large portion of the aggregate gender wage gap in the US. Similarly, Blau and Kahn (2016) show that human capital variables explained little of the observed gender gap in the US in 2010, in contrast with occupation and industry which account for over half of the gap. Fafchamps, Soderborn and Benhassine (2008) find that the sorting into occupations and firms explains the gender wage gap in a set of nine countries in Sub-Saharan Africa and two countries in North Africa.

We build on this work by examining the role of occupational sorting in gender wage gaps in Vietnam. We also explore some potential channels for emergence of occupational sorting there. Vietnam offers a great setting for these analyses - gender gaps in educational enrollment have been more or less closed and the rates of female labor force participation are relatively high (see Appendix I). From an analytical

perspective, this implies that differences in human capital and differential selection into the labor force are less likely to be big drivers of the gender wage gap. From a policy perspective, this setting supports a focus on occupational sorting in the task of tackling inequalities in economic opportunities between genders rather than on first order issues of labor force participation and education.

We start by showing that the gender wage gap in Vietnam persists despite the closing of the education gap and demonstrate that a large fraction of this wage gap is associated with occupational sorting. Then, we explore three hypotheses about how occupational sorting occurs:

- Occupational sorting occurs once women are in the labor market and is explained by sorting on the non-monetary characteristics of jobs. Women are more willing to forego monetary compensation for greater job security, insurance or leave since they have stronger preferences for these non-monetary characteristics, potentially due to non-equal distribution of care and household chores.
- Occupational sorting emerges much before women enter the labor market. Social norms about gender roles learned at an early age drive differences in career aspirations between young girls and boys. Career aspirations affect choices of the level of education and field of study, and thus predetermine occupational choices when girls enter the labor force.
- Sorting into different occupations occurs during school-to-work transition, because of gender specific barriers to finding jobs within one's field of study. We only test this hypothesis for individuals who have upper secondary, vocational or tertiary education, where specialization during study affects the types of jobs that are available.

We find support for the first hypothesis: our analysis suggests that women have stronger preferences for non-monetary job characteristics. We do not find evidence that differences in career aspirations of young boys and girls are a likely culprit for gender wage gaps in the future (the second hypothesis). Similarly, our results do not suggest that women face greater barriers in the school-to-work transition: they are not more likely compared to men to work outside of their field of study.

The remainder of this paper is structured as follows: in Section 2, we provide an overview of the relevant literature. In Sections 3 and 4, we present the empirical strategy and data sources respectively. We discuss the results in Section 5. Section 6 concludes.

2. Literature

This paper touches on several strands of literature: the dynamics of the gender wage gap in different parts of the world; the drivers behind the gender wage gap, including, the role of human capital vs. sorting; and finally, explanations of why occupation and industry sorting emerge.

Several studies have documented the decline in the gender wage gap in the US and OECD countries since at least the 1970s (Peterson and Morgan, 1995, Blau and Kahn, 2008). In developing countries, the trend is less clear-cut - Weichselbaumer and Winter-Ebmer (2005) find conflicting trends in different parts of the world in a meta-analysis. A common driver for the falling gap has been the narrowing of the gender gap in human capital, in terms of relative experience (O'Neill and Polacheck, 1993; Blau and Kahn, 1997) and in particular, education. Figure 1 shows that enrollment rates for boys and girls are similar in primary and secondary education in the most regions of the world. There is also reason to believe that increasing

enrollment among girls is translating into higher levels of educational attainment. Grant and Behrman (2010) show that conditional on ever being enrolled, girls have higher levels of educational attainment than boys, and across 38 less developed countries, girls progress through school at the same or faster pace than boys. Becker, Hubbard and Murphy (2010) show that the gender gap in college attainment had reversed in 67 of 120 countries in their sample between 1970 and 2010, including in countries with below-median per capita GDP.

Consequently, gaps in human capital now explain a smaller portion of the gender wage gap than in the past. Blau and Kahn (2016) show that human capital variables explained 27% of the gender wage gap in the US in 1980 but only about 8% in 2010. In contrast, occupation and industry choice explained 27% of the gap in 1980 and 49% in 2010. This is in line with findings from other countries – on average, human capital characteristics explain 5.3% of the wage gap in OECD countries and 6.4% in developing countries, but when controlling for occupation this reduces significantly to 3.6 and 3.2 % (Oostendorp 2009). What then explains occupation-industry sorting?

Sorting into industries has received less attention in the literature than sorting into occupations, but is important from a policy perspective. In the East Asian "miracle" economies, women were often clustered in the manufacturing sector, especially the export manufacturing sector. When Taiwan, China, experienced economic growth, manufacturing jobs were often lost to countries with lower labor costs; women in Taiwan, China, as a result enjoyed fewer of the benefits of that growth (Zveglich Jr and van der Meulen Rodgers, 2004).

Occupational sorting has been tackled in a few different ways in the literature. A first set of explanations focused on differences in levels of human capital (Mincer and Polachek, 1974). Specifically, greater human capital may raise productivity in some occupations more than in others. In this case, individuals are likely to sort across occupations based on their human capital level. However, as we note above, these differences have become less pronounced over time.

A second set of explanations focuses on differences in preferences and constraints between women and men. Currie and Chaykowski (1992) find differences in benefit coverage between male- and female-dominated occupations. Female dominated jobs had lower pension benefits but better paid- and unpaid-leave. They explain the variation using a model where women face a greater trade-off between time in household production and labor production.

A third set of explanations centers around gender differences in educational paths. Since many occupations today require occupation specific human capital that is acquired through education, gender streaming in education can have an effect on occupational choice and the gender pay gap (Black, Haviland, Sanders and Taylor, 2008). Women continue to lag in the STEM fields and particularly in mathematically-intensive fields (Ceci, Ginther, Kahn, and Williams, 2014). There is substantial evidence that mathematics test scores, math-based curricula and math as a college major are predictive of future income, while verbal abilities are not (Arcidiacono, 2004).

A plausible explanation of education streaming may be a performance gap between the genders that has been observed in different fields across different settings. In the US, girls and boys show no differences in

math or verbal skills at the start of education, but girls start to fall behind in math as early as fifth grade (Fryer and Levitt 2011). Furthermore, Niederle (2010) finds that boys are more likely to perform in the right tail of the distribution in the math section of the SAT, Advanced Placement tests (AP) and Graduate Record Examinations (GRE) in the US. Bharadwaj et al. (2012) find similar trends in middle- and low-income countries, such as Chile, Mexico and Thailand on the Programme for International Student Assessment (PISA) - around fourth grade, boys begin outperforming girls in math at both the median level and at top quantiles.

But there is considerable gender sorting that goes beyond performance. In the US and the Netherlands, boys are generally more likely to study science and math than girls even after controlling for performance, constructed from previous mathematics test scores (Buser, Niederle and Oosterbeek, 2014) or SAT scores (Turner and Bowen, 1999).[1] These gender specific choices which are not explained by performance may be due to the fact that girls have lower levels of self-assessment conditional on performance and that parents are more likely to invest in mathematics education for male children, as Bharadwaj et al. (2012) find in Chile. Vietnam, however, seems to be atypical with respect to these results. Dercon and Singh (2013) in a study of the Young Lives data from Ethiopia, India, Peru and Vietnam find that girls outperform boys in Vietnam and also that parents have a pro-girl bias in education aspirations for their children.

Another branch of literature has explored the role of psychological traits. This work suggests that women are more risk-averse and altruistic, have a lower preference for competition and are less willing to negotiate. Some of this evidence is context-specific. Women appear to exhibit more risk-aversion in their investments (Jianakoplos and Bernasek ,1998; Hinz, McCarthy and Turner, 1997), a pattern that persists even controlling for lower levels of income among women. Babcock et. al. (2017) show that women are more likely than men to be asked to volunteer to tasks with a low impact on the likelihood of a promotion and are also more likely to accept. Buser, Niederle and Oosterbeek (2014) find that a lower preference for competition, controlling for academic ability, explains about 20% of the gender difference in choice of academic track among high school students in the Netherlands. Bowles, Babcock and Lai (2006) demonstrate that there are differences between genders in the propensity to initiate negotiations in the US, with Card, Cardoso and Kline (2015) finding empirical support for the role of bargaining in explaining the gender earnings gap in the Portuguese labor market.

It is important to understand when these observed differences in these psychological traits emerge. For instance, Brinig (1995) empirically tests risk preferences with a sample of elementary, high school and graduate students and finds that gender alone does not explain differences in risk preferences, but it becomes a significant predictor when interacted with age. The author finds that gender differences in risk preference peak at the age of 30. In a field experiment in Israel, Gneezy and Rustichini (2004) find that boys react to competition, by running faster in races when competing against someone rather than running alone, while girls do not. Dreber, von Essen and Ranehill (2011), on the other hand, find that boys and girls react equally to competition in a similar experiment in Sweden. There is also mixed evidence in gender differences in risk taking at younger ages. Booth and Nolen (2012) find that among

[1] Buser, Niederle and Oosterbeek (2014) look at the track choices of Grade 9 students for their last three years of high school, whereas Turner and Bowen (1999) look at incoming university students' choice of college major.

10 and 11-year-old children, boys are more risk taking when compared to girls in mixed sex schools, but find no differences between boys and girls who attend single sex schools. Cardenas et al. (2012) compare competition and risk preferences among children ages 9-12 in Sweden, a country with a high gender equality index, and Colombia, a country with a lower gender equality index. In Colombia, girls and boys are equally competitive in four different tasks while in Sweden girls are more competitive than boys in certain tasks while boys are more competitive than girls in others. However, boys in both countries are more risk taking than girls, with the gap in risk taking being narrower in Sweden. While the evidence overall is varied, it does suggest that these preferences are not necessarily fixed at younger ages.

3. Empirical strategy

Our analysis consists of two stages. In the first, we empirically establish the premise of this paper - that the gender wage gap persists in Vietnam despite a narrowing of differences in levels of educational attainment, and that occupational sorting explains a non-trivial portion of this gap. In the second stage, we explore the three explanations for the emergence of occupational sorting. We describe each stage in more detail below.

a) Role of occupational sorting in gender wage gaps in Vietnam

To establish the persistence of the gender wage gap and the importance of occupational segregation, we rely on a decomposition approach similar to Peterson and Morgan (1995) and Fafchamps et al. (2009). This approach is purely descriptive. We attempt to document the magnitude of gender wage gap, and assess the importance of different factors, including education as well as occupational and industry sorting. The estimated coefficients on education or industry or occupation dummies can be interpreted as the return to education (or choice of occupation or industry) as well as its correlation with an unobserved ability term. We estimate the following equations:

$$LnY_i = \alpha + \beta_1 Female_i + \gamma X_i + \varepsilon_i \quad (1)$$

$$LnY_i = \alpha + \beta_2 Female_i + \gamma X_i + \Sigma_{p=1}^{P} \tau_p Edu_{ip} + \varepsilon_i \quad (2)$$

$$LnY_i = \alpha + \beta_3 Female_i + \gamma X_i + \Sigma_{p=1}^{P} \tau_p Edu_{ip} + \Sigma_{q=1}^{Q} \rho_q Occ_{iq} + \Sigma_{r=1}^{R} \omega_r Ind_{ir} + \varepsilon_i \quad (3)$$

where Y_i is our measure of earnings, X_i includes a set of controls for the province, whether the area is rural or urban, the respondent's age and whether he/she belongs to an ethnic minority, Edu_{ip} are dummies for the highest level of education, and Occ_{iq} and Ind_{ir} are dummies for the occupation and industry respectively.

The measure of the gender wage gap in each equation is β. Comparing the βs across the three equations provides a descriptive picture of relative contributions of education and choice of occupation and industry. For instance, comparing coefficients β_1 and β_2 reveals whether differences in the level of educational attainment between men and women explain some of the gender wage gap. If differences in education do not explain the gender wage gap, we would expect that $\beta_1 = \beta_2$. If on the other hand, the unconditional gender wage gap is driven by the fact that women have lower levels of educational attainment than men, we would expect that $\beta_1 > \beta_2$.

Similarly, comparing β_2 and β_3 reveals whether occupational segregation contributes to the gender wage gap. If men and women were equally likely to work in a specific industry and occupation, we would expect that controlling for occupation and industry would not change the observed wage gap ($\beta_2 = \beta_3$) . If on the other hand, following Peterson and Morgan (1995), we expect that women sort into lower paying occupations and industries, we would expect that $\beta_2 > \beta_3$.

b) Emergence of occupational and industry sorting

i. In the labor market: Sorting over non-monetary characteristics

Our first hypothesis is that occupational sorting is a result of sorting over the non-monetary characteristics of jobs. By non-monetary characteristics, we refer in this analysis to average weekly hours, having a formal job contract and having the following employment benefits: health insurance, social insurance and paid leave. Women may be more willing than men to forego monetary compensation in exchange for these non-monetary characteristics. We check if there is evidence that women who are currently employed as wage workers have a stronger preference for these non-monetary characteristics by estimating:

$$Prob(C_i) = \alpha + \beta_4 Female_i + \gamma X_i + \Sigma_{p=1}^{P} \tau_p Edu_{ip} + \varepsilon_i \quad (4)$$

$$H_i = \alpha + \beta_5 Female_i + \gamma X_i + \Sigma_{p=1}^{P} \tau_p Edu_{ip} + \varepsilon_i \quad (5)$$

where C_i is a dummy variable that takes a value of 1 if a specific non-monetary characteristic is available to individual i in their job, and H_i is the number of hours of work per week. We consider four non-monetary characteristics that are captured in the LFS: having a formal contract, health insurance, social insurance and paid leave.

To see whether these preferences contribute to the gender gap, we continue the decomposition exercise in $(1) - (3)$ adding controls for non-monetary characteristics. Specifically, we estimate:

$$LnY_i = \alpha + \beta_6 Female_i + \gamma X_i + \Sigma_{s=1}^{S} \mu_s C_{is} + \varepsilon_i \quad (6)$$

$$LnY_i = \alpha + \beta_7 Female_i + \gamma X_i + \Sigma_{p=1}^{P} \tau_p Edu_{ip} + \Sigma_{q=1}^{Q} \rho_q Occ_{iq} + \Sigma_{r=1}^{R} \omega_r Ind_{ir} + \Sigma_{s=1}^{S} \mu_s C_{is} + \varepsilon_i \quad (7)$$

Comparing β_6 to β_1 reveals the extent to which adding dummies for non-monetary characteristics C_{is} contributes to gender wage gap. If $\beta_6 < \beta_1$, one may conclude that a part of the observed gender wage gap is explained by gender specific sorting across non-monetary characteristics. Comparing β_7 to β_3 can tell us whether gender-specific differences in non-monetary characteristics of the jobs are fully explained by occupation-industry sorting, or whether women choose jobs with better non-monetary benefits within specific industries and occupations. $\beta_7 = \beta_3$would support the first argument, while $\beta_7 < \beta_3$ would support the second.

Taken together, results from (4) – (7) provide descriptive evidence of the role that gender specific preferences for non-monetary benefits play in occupational segregation and the gender wage gap. Equations (4) and (5) establish whether women in Vietnam are more likely to work in jobs with certain non-monetary characteristics. Equations (6) and (7) check whether sorting into jobs based on their non-monetary benefits contributes to the observed gender wage gap, and whether such sorting overlaps with sorting into specific industries and occupations.

ii. Before the labor market: The role of norms

Our second hypothesis is that streaming occurs well before men and women enter the labor market. This has important policy implications because some occupations and industries require occupation or industry specific human capital that needs to be accumulated starting at an early age. To become an engineer, for instance, one typically needs training at the tertiary level in the form of an engineering degree. But in order to pursue an engineering degree, one may need to have a minimum level of training in mathematics and science prior to entering university. This is affected by the choices one makes in choosing between the humanities and sciences tracks in high school, which may in turn be affected by a choice of which subjects to focus on in middle school. Consequently, decisions made at an early age may rule out specific occupations as career options at a later stage.

This is the case in Vietnam. Students must make choices that affect their career options early. Around age 15, they must decide on the field of study by choosing between exams in the social sciences and natural sciences, and between vocational and tertiary track. Those who choose vocational track, apply for technical training school, which may span from 6 months to 3 years. Those who choose tertiary track continue in school for grades 10 through 12, which allows to subsequently apply for a university.

We take advantage of existing data to probe the following question: do girls and boys aspire to different occupations as early as age 12? The threshold of 12 is chosen primarily due to data availability (discussed in the following section), but it coincides with the age at which gender differences in performance and psychological traits have been observed in economic literature (see for instance Fryer and Levitt, 2011). The literature in psychology suggests that orientation to sex roles forms as early as between the ages of 6 and 8[2] (Gottfredson, 1981, Bian, Leslie and Cimpian, 2017, Cvencek, Meltzoff and Greenwald, 2011). Based on these literatures, it seems reasonable to expect to see differences in aspirations at age 12 if our hypothesis is correct. Moreover, there is relatively little economic literature on the question of whether gender specific labor market preferences are shaped years before entering the labor market. Our work will be a first attempt to fill this gap,[3] and further work can address the question of at what age these preferences first emerge.

We map the occupations that the 12-year-old boys and girls in our data aspire to, to the average monetary and non-monetary characteristics of these occupations in the concurrent labor force survey data. We construct the aspirations variable from the following survey question: “When you are about 25 years old, what job would you like to be doing?” Combining the two surveys allows us to answer the following question: if boys and girls succeed in pursuing occupations they aspire to at the age of 12, and the relative wages observed at the time of the survey do not change, would we observe a gender wage gap? It is important to note that we do not assume that boys and girls at 12 make choices based on expected earnings or non-monetary characteristics. Rather, we believe that at an early age these aspirational choices are made based on idealized perceptions of jobs, where social norms play an important role. Social

[2] Between the ages of 5 and 6, girls change from associating their own gender with “really, really smart” to associating boys with “really, really smart,” whereas boys change from associating their own gender with “really, really nice” to associating girls with “really, really nice (Bian, Leslie and Cimpian, 2017).

[3] For literature in psychology and sociology, see for example Oosterbeek (2014) and Gneezy and Rustichini (2004).

norms affect perceptions about which occupations are seen as interesting and exciting, as well as which are suitable for men and women.

Consequently, we expect that if (i) social norms mandate that some occupations are more suitable for men than for women, (ii) these occupations tend to be better remunerated, and (iii) if social norms play an important role in aspirations of the current generation, we will observe a gender wage gap in a simulated distribution of earnings, based on stated aspirations for future jobs, and median earnings that correspond to these jobs in the current LFS.

We specifically check this by estimating the following simple regression:

$$LnY_i = \alpha + \beta_8 Female_i + \epsilon_i \quad (8)$$

where Y_i is a measure of the aspirational wage, which we define as the median wage in the occupation that a respondent i aspires to. β_8 significantly less than 0 suggests that girls already aspire to lower paying occupations than boys by the age of 12. We also test to what extent gender gap in aspirational earnings varies depending on the level of ability by estimating:

$$LnY_i = \alpha + \beta_9 Female_i + \gamma\theta_i + \varepsilon_i \quad (9)$$

$$LnY_i = \alpha + \beta_{10} Female_i + \gamma\theta_i + \rho(Female_i * \theta_i) + \varepsilon_i \quad (10)$$

where θ is a measure of ability. Equation (9) captures the gender gap in aspirational wages conditional on ability, and equation (10) allows for variation in the size of the gender gap at different levels of θ i.e., if the coefficients β_{10} and ρ are both negative and significant in (10), the gender wage gap increases at higher levels of ability. Understanding how aspirational gap varies along the ability spectrum could further provide useful information on targeting potential interventions.

We also test whether we would observe gender gaps in non-monetary characteristics of jobs if boys and girls succeeded in pursuing their aspirations and relative non-monetary characteristics of the jobs remained the same. To do so, we once again map aspirational occupations to concurrent labor force surveys, focusing on non-monetary characteristics. The approach is identical to that in $(8) - (10)$, but using P_i - a proportion of jobs within each occupation that have a specific non-monetary characteristic - as the dependent variable. The set of non-monetary characteristics remains the same as in Section b.i.: having a formal contract, health insurance, social insurance and paid leave, and average weekly hours. As before, we take advantage of availability of data on ability to explore whether preferences for non-monetary benefits vary depending on the ability level and how ability interacts with gender:

$$P_i = \alpha + \beta_9 Female_i + \varepsilon_i \quad (11)$$

$$P_i = \alpha + \beta_{10} Female_i + \gamma Ability_i + \varepsilon_i \quad (12)$$

$$P_i = \alpha + \beta_{11} Female_i + \gamma Ability_i + \rho(Female_i * Ability_i) + \varepsilon_i \quad (13)$$

iii. Transition to the labor market

Even if women choose the same fields of study as men, we may observe occupational sorting if there are gender specific barriers to finding jobs within one's field of study. In this case, occupational sorting occurs not at a young age when boys and girls make decisions about what type of human capital to accumulate

(hypothesis 2), nor once they are in the labor force and have to make a trade-off between monetary and non-monetary benefits (hypothesis 1), but during the transition to the labor market. This scenario would require a different set of policies.

To test this hypothesis, we check whether women are more likely than men to be working outside their field of study. We carry out a matching exercise between field of study and occupation to identify whether a worker is working within their field of study or working outside of their field of study. Specifically, we estimate:

$$Pr(Mismatch) = \alpha + \beta_{14} Female_i + \gamma X_i + \varepsilon_i \quad (14)$$

where $Mismatch$ is a dummy variable that takes the value of 1 when person i works outside their field of study, and 0 otherwise. A positive β_{14} would support the hypothesis that women find it more difficult than men to find jobs within their field of study. X_i here only includes a control for age.

4. Data

We use three data sets to implement the strategy described above. First, we use data on wage workers in Vietnam from the Labor Force Surveys (LFS) carried out by the General Statistics Office (GSO) of Vietnam between 2011 and 2014. The LFS includes information on occupation, industry of employment and highest level of educational attainment for all respondents. Occupations are classified using the international standard classification of occupations (ISCO-08) and industries using the International standard classification of Industry (ISIC-08), with both available at the 4-digit level (corresponding to more than 400 occupations and industries), providing a sufficiently granular level of information on occupational and industry choice. A worker's highest level of education is classified using a system that is based on the international standard classification of education (ICSED-08). Vietnam has an extensive system of vocational education, so we simplify the educational classification for ease of presentation[4] (see Appendix III).

The LFS data permits us to estimate equations 1 - 3 – i.e. to estimate the size of the gender wage gap in Vietnam, the role that differences in educational attainment and occupation-industry sorting play in explaining the size of the gender wage gap. Given the sample design of the LFS, these estimates likely provide a good picture of the overall state of the labor market in Vietnam.

Second, we use data from the Vietnam Young Lives project. Young Lives is an international longitudinal cohort study of childhood poverty which has been following 12,000 children in four countries (Ethiopia, Peru, Vietnam and India) over a period of 15 years. Each country has two cohorts, one consisting of 2,000 children born in 2001-2002 and another consisting of 1,000 children born in 1994-1995. The survey consists of three main elements: (1) a household questionnaire, (2) a child questionnaire and (3) a community questionnaire.

The first of four rounds of data collection under Young Lives (YL) took place in 2002. It used a sampling strategy designed to ensure over-sampling of poor communities. An equal number of eligible children

[4] We originally ran the analysis on the fully extended system of educational levels. The results are robust to the choice of classification system.

(aged either 1 or 8) were randomly selected into the sample in each community. Three more rounds of data collection took place in 2006 (Round 2), 2009 (Round 3) and 2013 (Round 4) – when the children were aged between 11 and 12 years (the Younger Cohort) and between 18 and 19 years (the Older Cohort). We use data from Round 4 for Younger Cohort. We prefer these data for two reasons. First, round 4 of data collection falls within the time-frame of the LFS data we use for the study, and second, due to differences in the questionnaires administered to the older and younger cohort, fine measures of cognitive ability are only available for the younger cohort in this round.

We use data from the child questionnaire to explore gender differences in aspirations for a future job, controlling for cognitive ability. The Young Lives project used the Raven's Matrix test in the first round and a Peabody Picture Vocabulary test in the following rounds, using Itemized Response Theory (IRT) in subsequent rounds to provide a standardized and comparable measure of cognitive ability across the rounds. For the 4th round, the PPVT test was only administered to the younger cohort; the older cohort was given math and vocabulary tests.

We match the occupation that children claim they would like to have when they are 25 years old to occupations in the 2013 LFS (same year as Round 4) so that we can create the monetary and non-monetary characteristics associated with these aspirational occupations based on the LFS data. On-line Appendix provides more details on the matching procedure.

Our final source of data allows us to examine the relationship between field of study and occupation, and comes from the World Bank's Skills Towards Employability and Productivity (STEP) program. The STEP program is a World Bank initiative to measure the relationship between skills, employability and productivity in low- and middle-income countries. STEP targets the urban working age population, defined as all adults aged 15-64. In Vietnam, the surveys were carried out in Ha Noi and Ho Chi Minh City, with households selected through a process of stratified random selection and the primary respondent within each household chosen at random. We use data from the first wave of STEP surveys carried out by the World Bank in Vietnam in 2012, since this falls within the time period covered by our LFS data.

STEP includes an employee module and an employer module and covers 3,405 adults, of whom 2,037 are women. We use data from the employee module and restrict our analysis to adults who are currently employed as employees (rather than employers or self-employed) to make the sample comparable to the pool of wageworkers from the LFS. This reduces the sample to 1,316 observations (643 women).

The STEP data have information on occupation and industry at the 3-digit level, which is one level of aggregation up from the 4-digit classification available in LFS. Nevertheless, this corresponds to over 100 occupations and industries in our data. STEP makes a distinction between general and vocational education, using a different survey question to capture the level of educational attainment in each. We combine the responses to these questions and simplify the levels of response to make them similar to the ICSED-08 classification system used in the LFS. This leaves us with 8 levels of education (details of mapping are presented in Appendix III).

5. Findings

In this section, we present the results from the estimation of the regressions presented in Section 3 and using the data described in Section 4. Our preferred earnings measure is hourly total compensation, which includes wage income and any non-wage payments such as overtime benefits, measured in nominal terms.[5] As a robustness check we use hourly wages instead of hourly total compensation, and our results do not change.[6] An hourly measure is preferred since there are differences in hours worked between men and women. The top and bottom 0.1% of earnings are trimmed to avoid the outsized influence of outliers. We estimate equations (4) and (14) using a linear probability model for ease of interpretation, and the remaining equations using an OLS regression and robust standard errors.

a) Gender wage gap and occupational sorting

We estimate equation (1) using the data from the Vietnam LFS.[7] The results suggest that women earned consistently less than men[8] on average across the four rounds of the LFS survey that we included in our study. The coefficient on the female dummy (β_1) is negative and significant in the earnings function in each of the four years. As figure 3 demonstrates, the gap between male and female earnings remains consistent – corresponding in terms of annual earnings to about 3,000,000 Dong, or $130.[9] However, the hourly compensation increased substantially over this period, and as a result, the gap falls in relative terms from close to 10% in 2011 to nearer 5% in 2014. Note that this is smaller than the gender wage gap observed in other countries – for instance, Blau and Kahn (2016) find that the ratio of unadjusted female earnings to male earnings in the US in 2010 was 79% and adjusting for covariates including human capital and occupation-industry increases this figure to 91%.

[5] We do not adjust for inflation since the focus of our analysis is on within-year comparisons of female earnings to male earnings rather than across year comparisons.

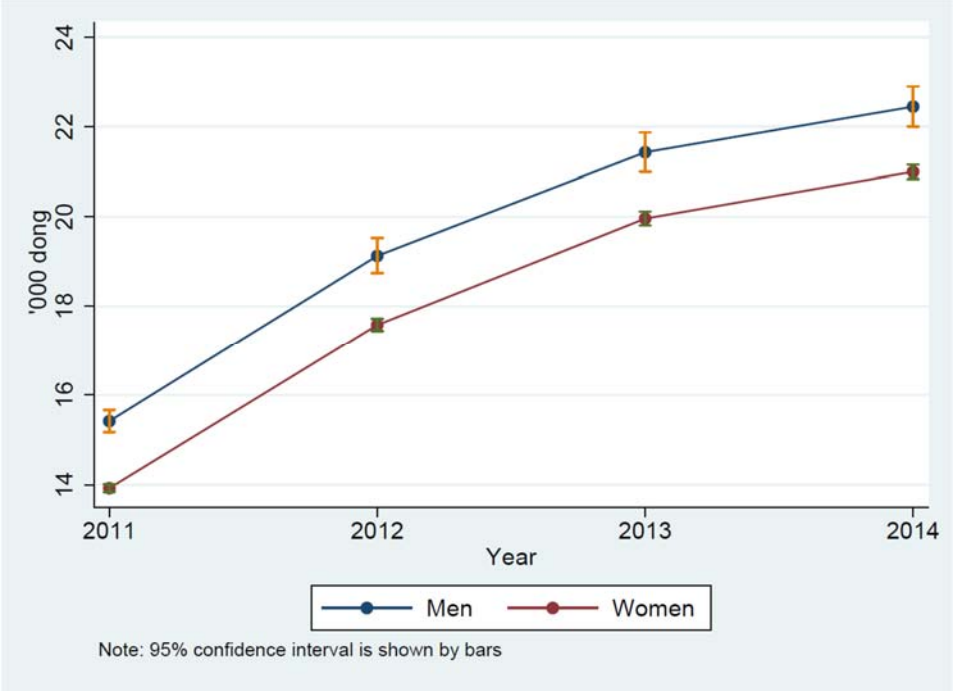
[6] These results are available from the authors upon request.

[7] Appendix II presents descriptive statistics of the sample.

[8] Earnings data in the LFS, and typically in other labor force surveys, are confined to respondents who are employees or "wage workers", and therefore do not include earnings of own-account, self-employed or unpaid workers. It is worth noting that in Vietnam, women are less likely to be in the labor force or be wage workers and more likely to be unpaid workers (see Appendix II for more). The gap in earnings that we examine therefore represents only a partial picture of the difference in economic opportunities between the genders in Vietnam.

[9] Assuming a 40-hour work week and 52 working weeks. Conversion to USD based on an exchange rate in 2017 of 22,700 dong = 1 USD.

Figure 3: Total (nominal) hourly compensation for men and women in Vietanm (from Labour Force Surveys).

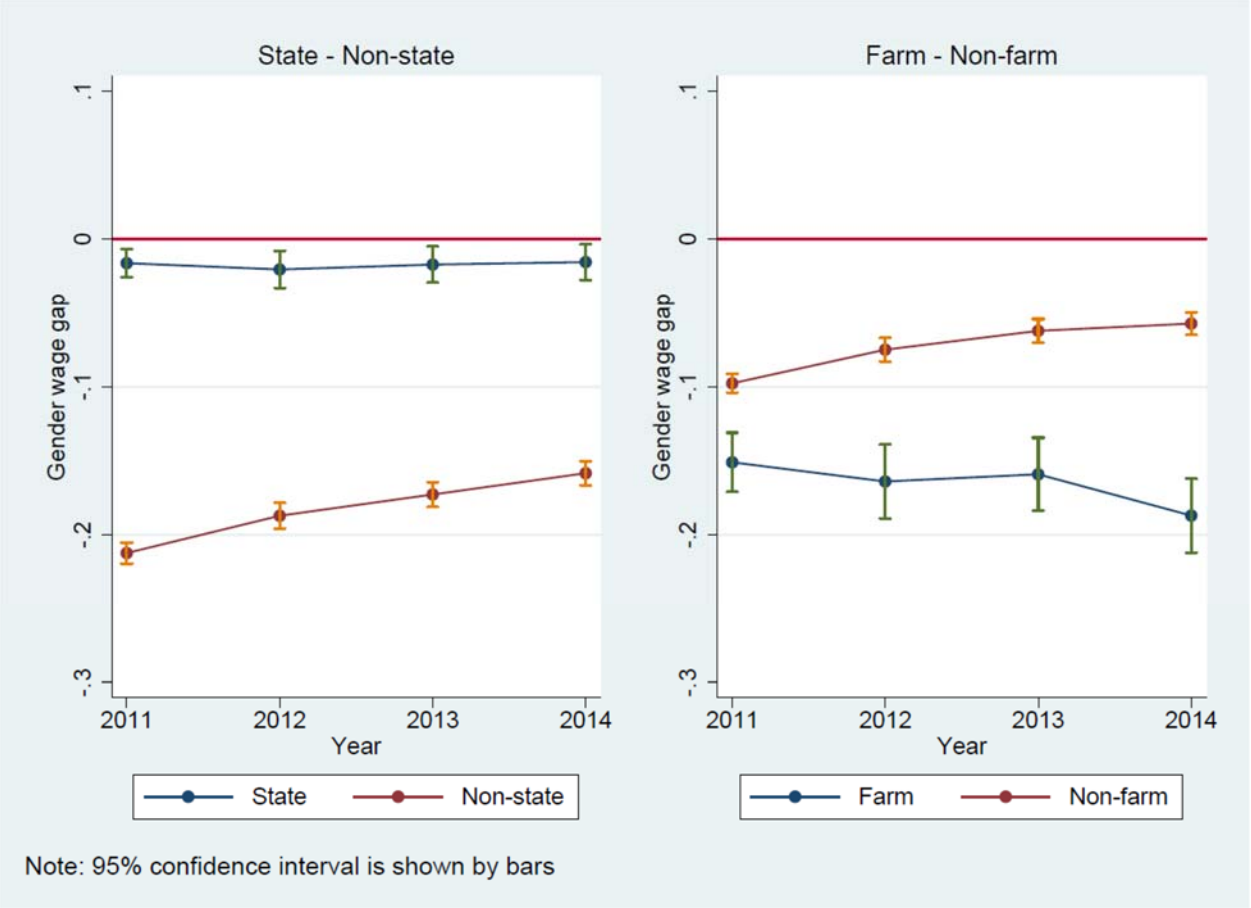


When we examine the size of the gender wage gap across different age groups, we see that the earnings gap is more or less similar in magnitude, with the notable exception of the cohort aged 55-59. This age group includes the retirement age for women in Vietnam: 55. It may be that women who continue working beyond the retirement age may be concentrated in lower paying occupations, or that women switch to lower paying occupations at retirement. This is consistent with the fact that the gap narrows once again in the 60+ cohort, when men approach retirement age (60) and become subject to the same labor market constraints.

Figure 4: Gender wage gaps in each age cohort

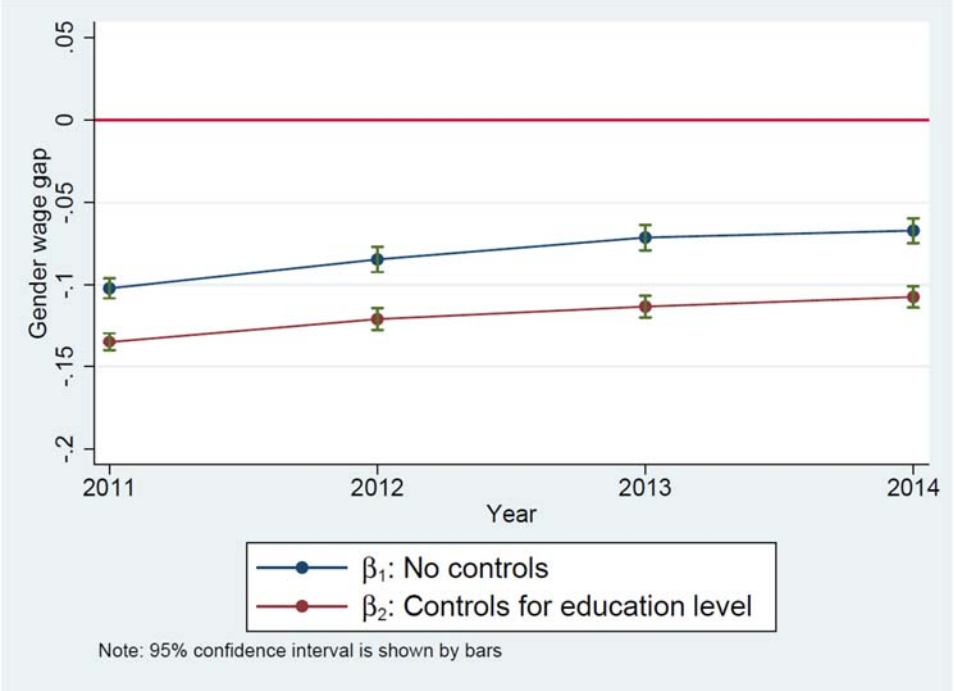
Next, we analyze the earnings gap within sectors and types of establishments to see if the gap is driven by a subsample of the workforce. We estimate equation (1) on subsamples of workers who work in state enterprises and in non-state enterprises, and those engaged in farm and non-farm work. We observe that the gap is present in both state and non-state enterprises, although the gap is much larger in the non-state enterprises (Figure 5). Better enforcement of labor market protections or less discretionary pay in state agencies or state-owned enterprises may explain this result. The gaps are also much larger in the agricultural sector than in the non-farm sector, perhaps as a result of the relatively high level of occupational segregation in agriculture in South East Asia (World Bank, 2008).

Figure 5: Gender wage gaps in different subsamples in VIetnam



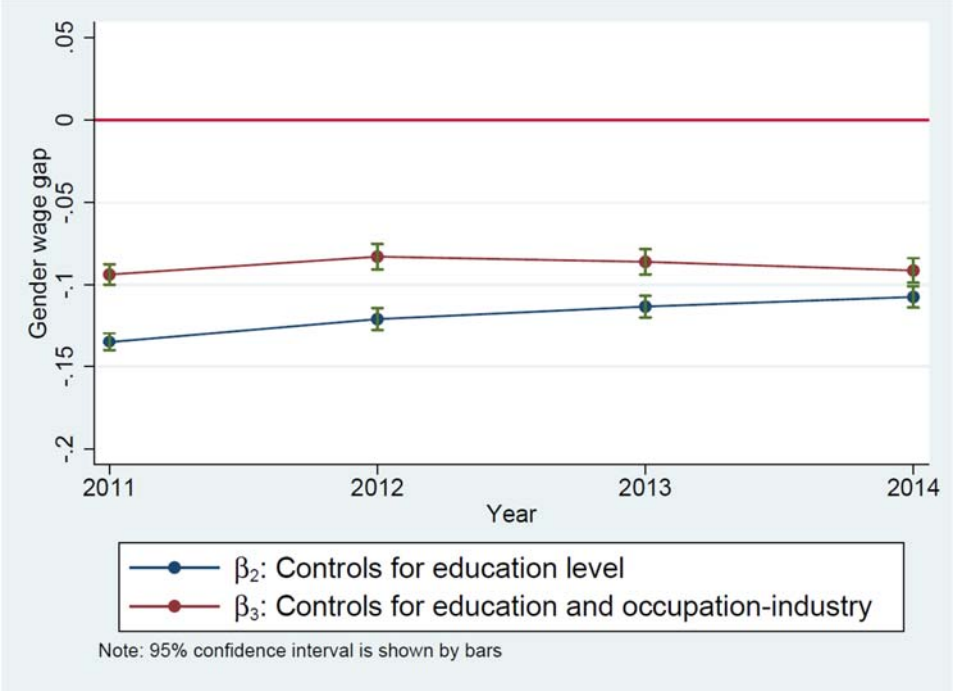
Our primary takeaway from these findings is that while there is heterogeneity in the magnitude of the gap, there is consistently an earnings gap between men and women in our data. The premise of this paper is that this gap persists even though women have closed the gap in education levels; and one of the key factors driving this gap is women sorting into lower paying occupations. To verify this claim, we present the results from estimating equations (1) – (3). If the earnings gap is driven by women having lower levels of educational attainment, we would expect that the gap narrows when controlling for education (i.e. $\beta_1 > \beta_2$). Figure 6 plots the estimates of β_1 and β_2, and suggests that in fact the opposite is true: in the presence of controls for education, the earnings gap between men and women widens. In fact, it is striking that the magnitude of the gender wage gap conditional on the level of education is 50% larger.

Figure 6: Gender wage gap in the presence of controls for education levels.



Similarly, comparing β_2 and β_3 sheds light on whether segregation across occupations and industries contributes to the gender wage gap. If men and women were equally likely to work in a specific industry and occupation, we would expect that controlling for occupation and industry would not change the observed gender wage gap ($\beta_2 = \beta_3$) . However, like Peterson and Morgan (1995), we find that the gender wage gap narrows when we control for occupation and industry ($\beta_2 > \beta_3$) (Figure 7). This result is in line with evidence from the US and OECD countries (Blau and Kahn, 2016; Oostendorp, 2009). There is some heterogeneity at different levels of education; specifically, occupational segregation seems to matter the most at the lower levels of education. However, overall occupational segregation appears to be a key contributor to the gender wage gap in Vietnam.

Figure 7: Gender wage gap in the presence of occupation-industry controls.



b) The emergence of occupational sorting

Having established that the gender wage gap persists in Vietnam despite women closing the education gap, and the role of occupation-industry sorting, next we evaluate our three hypotheses about the emergence of occupational sorting.

i. Sorting over non-monetary job characteristics

Our first hypothesis is that sorting over non-monetary characteristics explains the greater likelihood of women working in lower paying occupations. Women may have stronger preferences for having a formal contract, health insurance, social insurance and paid leave, and forego monetary compensation for these benefits. We examine whether this is the case in the labor market in Vietnam among those currently employed as wage workers, by estimating equations (4) and (5). Figure 8 shows that women[10] indeed are more likely than men to be working in jobs where they have a formal contract, paid leave, health insurance and social insurance. Figure 9 shows that women also work fewer hours per week on average than men.

[10] Note that since our sample here is wage workers covered in the LFS, these results must be treated as being true conditional on being a wage worker in Vietnam.

Figure 8: Likelihood of having each non-monetary characteristic, relative to men.

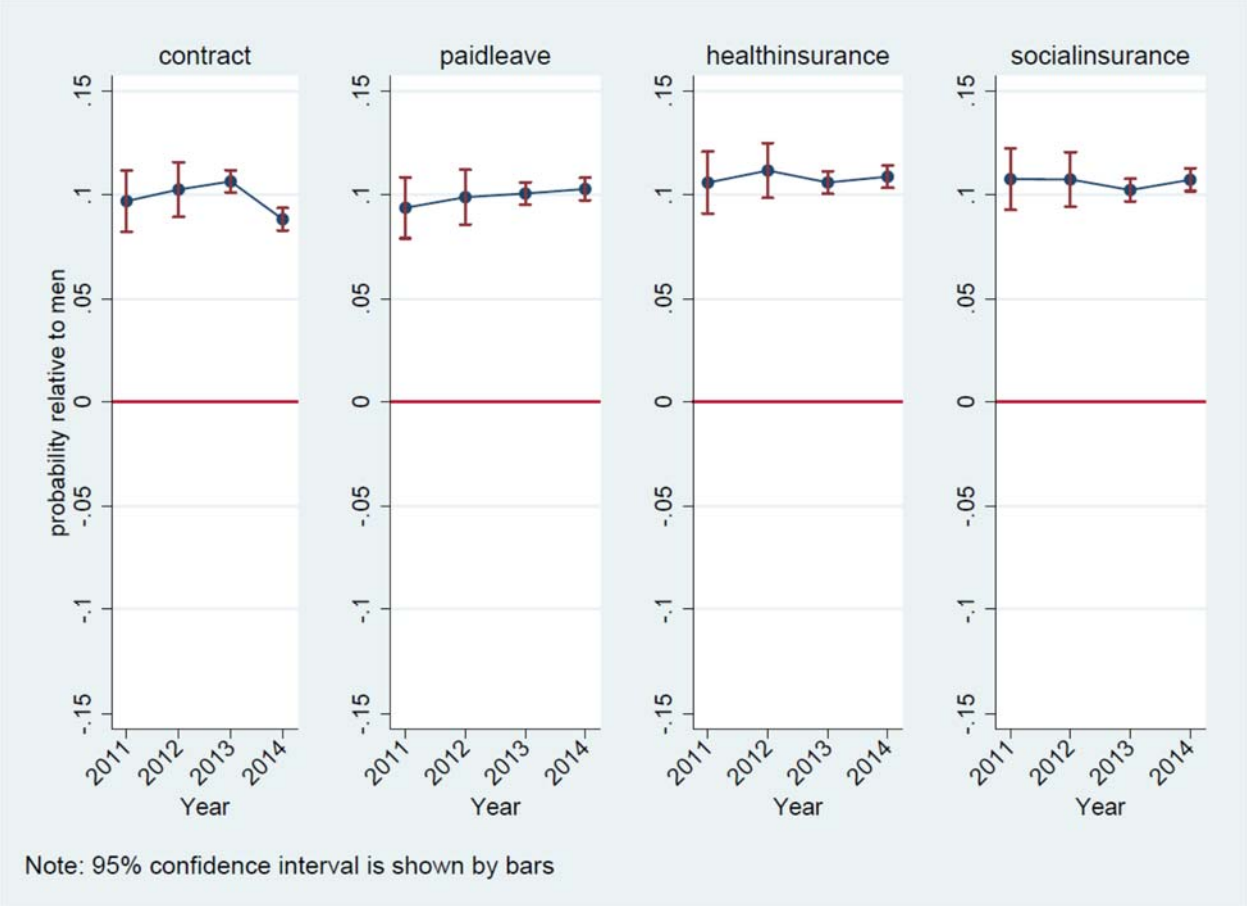
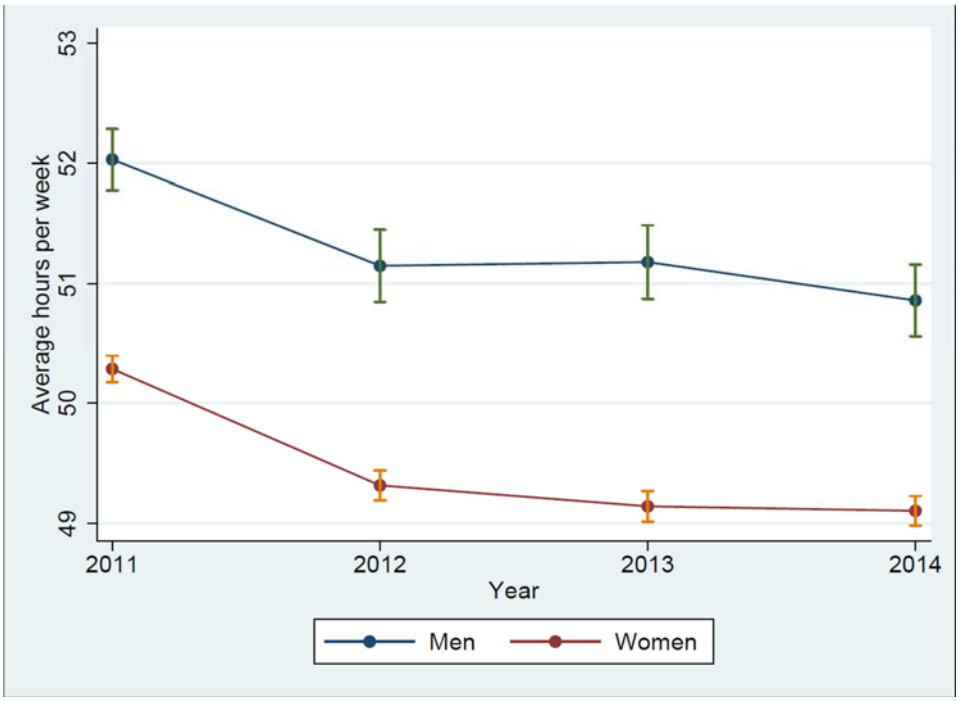


Figure 9: Hours worked per week.



Taking the results from figures 8 and 9 to indicate a stronger preference among women for favorable non-monetary job characteristics, figures 10 and 11 help us understand to what extent sorting over these preferences overlaps with occupational sorting. First, Figure 11 compares the gender wage gap from equation (1) with the gender wage gap from equation (6) and shows that adding controls for non-monetary characteristics widens the gender wage gap. This implies that there is a large gender wage gap even within jobs that share similar non-monetary characteristics. Figure 12 shows that adding controls for occupation and industry reduces the size of the gap substantially (β_7). The size of the gap however, is not significantly different from the size of the gap in the presence of occupation-industry controls alone (β_3). This suggests that variations in non-monetary characteristics within occupation and industry do not explain the gender wage gap. Rather, women sort into occupations and industries that offer a better package of non-monetary characteristics.

Figure 10: Gender wage gap widens when controlling for non-monetary characteristics.

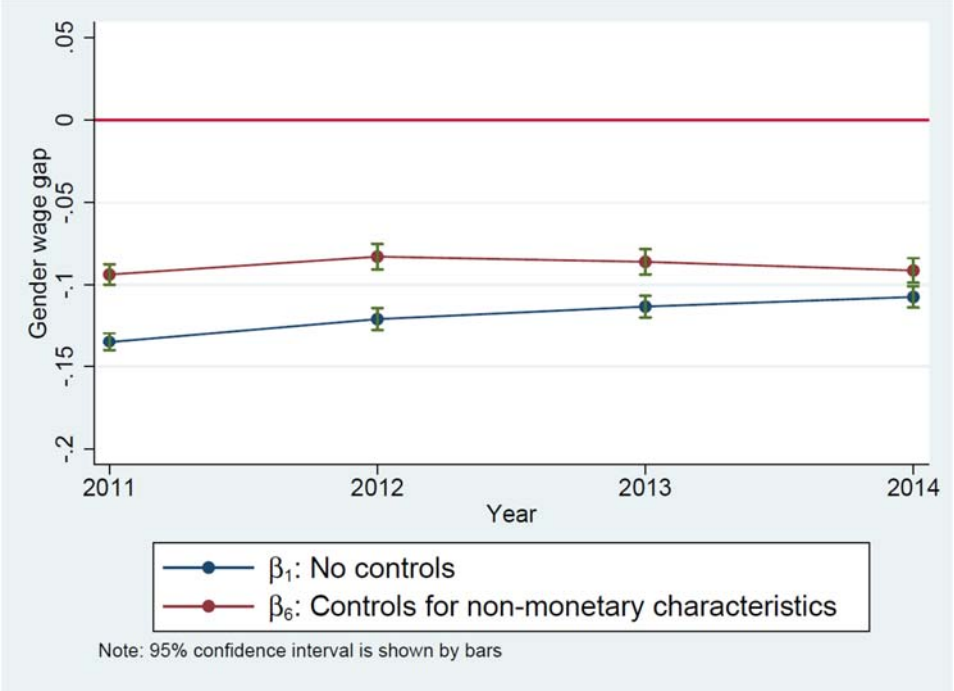
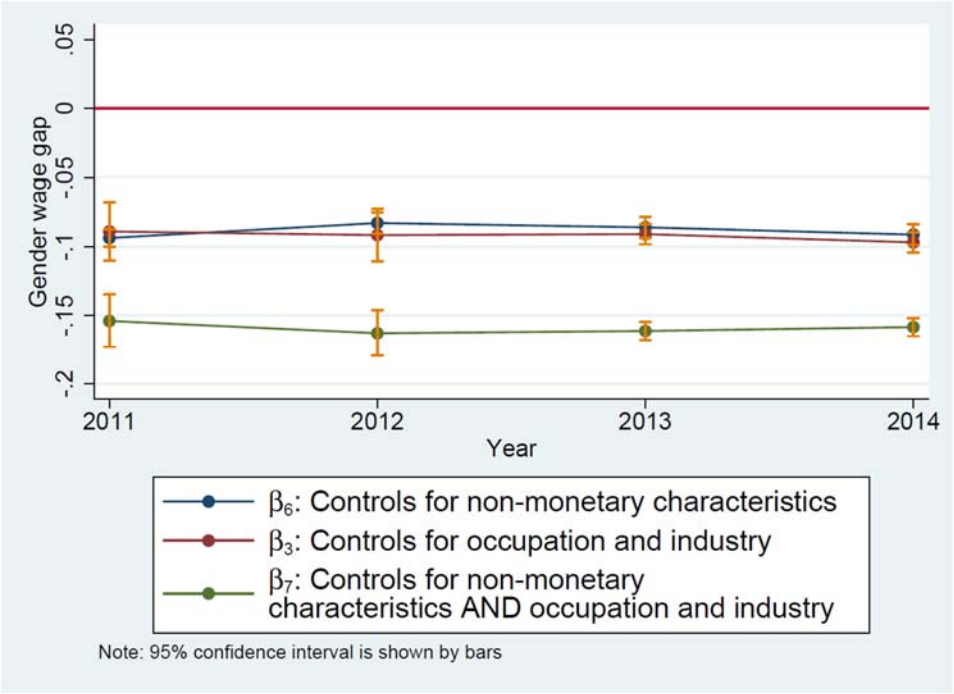


Figure 11: Gender wage gap with both non-monetary characteristics and occupation-industry controls is similar to gender gap with occupation-indystry controls alone.



ii. Social Norms

Next, we ask if gender streaming takes root at an early age through the effects of social norms. As described above, we use data on the aspirations of 12-year-old children and map the jobs that these

children want to have at the age of 25 to the data on wages and other characteristics of these jobs in the LFS for the same year. Using a natural log of median wage for the occupation as a dependent variable, we estimate equations (8) – (10). As we highlight in the discussion of our empirical strategy, the idea behind this line of questioning is not to claim that the aspirations of boys and girls at age 12 are based on having perfect information on earnings and the non-monetary characteristics of jobs. It is rather to ask a tangential question: *if* boys and girls made education and career choices based on their aspirations at the age of 12, would we still observe a gender wage gap?

As the results in Table 1 show, there is a reverse gender gap in aspirations among the 12-year-olds in our study - at age 12 girls aspire to jobs that have higher median wages than boys, even after controlling for levels of ability. These results suggest that if girls and boys aged 12 were to pursue and realize their career aspirations and the relative wages for different occupations remained the same, female earnings would on average exceed male earnings. Interestingly, the coefficient on the interaction term between the female dummy and the measure of cognitive ability is negative, although smaller in magnitude than the coefficient on ability. Although girls aspire to higher paying jobs than boys (and girls with higher levels of measured cognitive ability aspire to higher paying jobs), the difference in aspirations between the genders becomes smaller at higher levels of ability.

Table 1: gender wage gap, based on aspirational occupations

	Log Median Wage			
	(1)	(2)	(3)	(4)
Female	0.079***	0.080***	0.263***	0.256***
	(0.012)	(0.012)	(0.091)	(0.091)
PPVT		0.002**	0.003***	0.003**
		(0.001)	(0.001)	(0.001)
Female*PPVT			-0.003**	-0.003**
			(0.002)	(0.002)
Other controls				*Yes*
R2	0.035	0.038	0.042	0.075
N	1,180	1,180	1,180	1,180

The absence of a gender wage gap, extrapolated from the aspirational occupations of boys and girls at age 12, at first glance appears at odds with the abundance of informal evidence on the prevalence of norms and gender stereotyping in most societies. However, one mitigating factor here may be that the extent of stereotyping varies between societies and may change overtime. Further research is needed to explore this phenomenon in Vietnam.

How is a gender gap in aspirational earnings favoring girls at the age of 12 compatible with lower earnings for women as adults in Vietnam? A few explanations seem plausible and we test two in this paper. Girls may make human capital accumulation decisions similar to boys, so that streaming does not occur before the labor market. When women enter the labor market, however, they may face difficulties in finding jobs within their field of study. We explore this hypothesis in the subsequent section. On the other hand,

women may choose to forego higher earnings for a better package of non-monetary benefits. Section b) i., indeed, demonstrates some evidence in support of this hypothesis.

As discussed earlier, such preference for a better non-monetary package is likely to be driven by the fact that in Vietnam, as in most societies, social norms place the burden of care and household responsibilities in the purview of women. Notably, although the *need* to balance market and household roles may only emerge in adulthood, it is worth noting that the Young Lives data suggest that already at the age of 12 girls seem to also aspire to occupations with better non-monetary characteristics, in addition to occupations with higher wages. Table 2 shows that girls aspire to occupations where the non-monetary characteristics we analyze in this paper are more common - formal contract (Columns 1-4), paid leave (Columns 5-8), health insurance (9-12), social insurance (13-16). Table 3 shows preference for occupations with shorter weekly hours. The gender gap seems to narrow at higher levels of ability for most characteristics – the coefficient on the interaction term is significant in Columns 3, 9 and 12 of Table 2.

Table 2: Preferences for non-monetary characteristics among 12-year-old girls and boys.

	Formal Contract				Paid Leave			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.054***	0.054***	0.158***	0.140***	0.153***	0.154***	0.297***	0.284***
	(0.006)	(0.006)	(0.046)	(0.046)	(0.013)	(0.013)	(0.097)	(0.097)
PPVT		0.002***	0.003***	0.002***		0.002***	0.003***	0.003**
		(0.000)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Female*PPVT			-0.002**	-0.001*			-0.002	-0.002
			(0.001)	(0.001)			(0.002)	(0.002)
Other controls				*Yes*				*Yes*
R2	0.060	0.087	0.091	0.123	0.106	0.112	0.113	0.134
N	1,180	1,180	1,180	1,180	1,180	1,180	1,180	1,180

	Health Insurance				Social Insurance			
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Female	0.198***	0.199***	0.403***	0.378***	0.199***	0.199***	0.407***	0.383***
	(0.016)	(0.016)	(0.120)	(0.120)	(0.016)	(0.016)	(0.120)	(0.120)
PPVT		0.003***	0.005***	0.004**		0.003***	0.005***	0.004**
		(0.001)	(0.001)	(0.002)		(0.001)	(0.001)	(0.002)
Female*PPVT			-0.003*	-0.003			-0.004*	-0.003
			(0.002)	(0.002)			(0.002)	(0.002)
Other controls				*Yes*				*Yes*
R2	0.114	0.123	0.125	0.147	0.114	0.123	0.125	0.147
N	1,180	1,180	1,180	1,180	1,180	1,180	1,180	1,180

Table 3: Average weekly hours of jobs aspired to at age 12.

	Average hours per week			
	(1)	(2)	(3)	(4)
Female	-2.098***	-2.089***	-1.205	-1.491
	(0.155)	(0.154)	(1.153)	(1.150)
PPVT		0.035***	0.043***	0.011
		(0.010)	(0.014)	(0.015)
Female*PPVT			-0.015	-0.010
			(0.019)	(0.019)
Other controls				*Yes*
R2	0.134	0.144	0.144	0.180
N	1,180	1,180	1,180	1,180

It is possible that social norms governing career aspirations may become stronger after the age of 12 but before entering the labor market. This would be consistent with the evidence from several studies that find that gender differences in performance and psychological traits begin to emerge around the age of 10-12 and widen thereafter. To confirm that this is the case, we would repeat this type of inquiry for girls and boys aged 15 or 16, if the data become available for future rounds of the Young Lives survey. The analysis of the STEP data in the next section provides some initial support for this argument. In addition to any barriers in the school-to-work transition, there appears to be some streaming before the labor market – women are overrepresented in lower paying fields of study compared to men (Figure 12).

One explanation that does not seem to be backed by the data is that social norms may be changing in Vietnam. Although we cannot rule out the possibility that the female YL respondents will enjoy higher salaries than male YL respondents when they eventually enter the labor market, analysis of the change in the gender wage gaps across different cohorts suggests that this is unlikely. Figure 5 suggests that gender gaps are more or less at the same level across different cohorts and do not show signs of having narrowed in younger cohorts, which would be an indication of changing norms.

It is also worth considering some of the limitations of our analysis. The choice of future occupations in the YL data was limited to a set that is much narrower than the options reported in the LFS data. Consequently, when interpreting results, we should keep in mind that the children were choosing from a limited number of options. Nonetheless, the results suggest that at the age of 12 and at least in aspirations girls are not constrained by restrictive social norms about what jobs are appropriate for men and what jobs are appropriate for women.

iii. Labor market transition

Our third hypothesis about the emergence of streaming is that streaming resulting in gender wage gaps does not occur prior to the labor market, but happens during the school-to-work transition. Gender differences in the school-to-work transition could occur in the presence of gender-specific barriers to entry for some fields. One indication of such differences between genders would be a higher fraction of women working outside of their field of study. We use data on field-of-study available in the World Bank's STEP survey to check whether we observe a mismatch between field of study and occupation more often

among women in the STEP sample. Notably, in this case gender gaps due to streaming would be exacerbated by the education-occupation mismatch penalty - Nordin et al. (2010) and Robst (2007a) show that such mismatch is associated with a negative effect on wages.

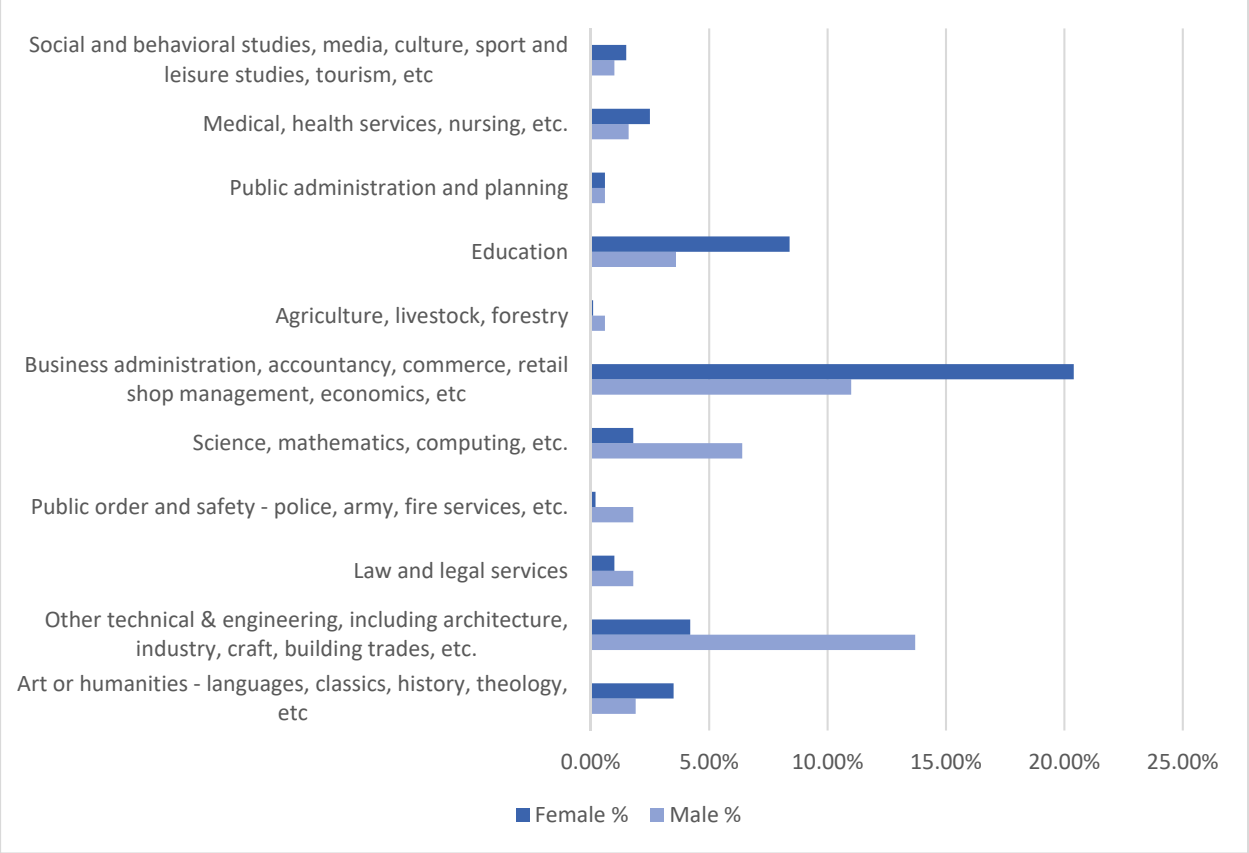
In order to construct the $Mismatch_i$ dummy which indicates a mismatch between education and occupation, we match occupations to fields of study using our own classification (presented in the Online Appendix), based on the classification in Montt (2015). Since specialization in the field of study tends to occur at higher levels of education, we restrict the sample to respondents who have completed the equivalent of upper-secondary education (including vocational school).[11] Interestingly, in addition to asking respondents about their current occupation, STEP also asks about the respondent's first occupation. This allows us to check if women are mismatched when they first enter the labor market or switch into mismatched occupations subsequently.

We first look for signs of streaming in education in the STEP data. Ranking fields of study by the median wages,[12] we find that the proportion of women in fields at the lower end of the earnings scale like "Education" exceeds the proportion of the men in those fields. The opposite is true at the higher end where the proportion of men in fields like "Science, mathematics and computing" or "Other technical & engineering fields" exceeds the proportion of the women. The exception to this trend is "Arts or humanities", which appears to be the field with the highest median wage in our sample – here more women are in the field than men. Overall, Figure 12 provides suggestive evidence of some streaming in education in Vietnam.

[11] Specifically, we include respondents with Upper Secondary or University education (corresponding to a level of education equivalent to ICSED 4 or higher) and respondents who have attended vocational school/trade school or vocational college.

[12] We only consider the wages of individuals who work in an occupation that matches their field of study, since this arguably provides a better estimate of the expected earnings of a field.

*Figure 12: Proportion of Male and Female populations in each field of study, in order of increasing median wage**



** "General" and "Other" fields excluded from figure, hence the percentages do not sum up to 100.*

Taken together with our finding that there is no gender wage gap in aspirations at age 12, Figure 12 allows for narrowing down the period when educational streaming may emerge: after the age of 12, but prior to graduation from high school. However, the effects of educational streaming can be exacerbated or mitigated during the transition to the labor market in the presence of gender specific barriers.[13] We estimate equation (14) to look for whether such differences lead to a great degree of education-occupation mismatch between female and male workers. Table 4 presents the likelihood of an education-occupation mismatch for women, relative to men, both in current occupation and first occupation.

[13] One example of such barriers is captured in the way men and women search for jobs. STEP asks men and women how they carry out their job search and Figure 16 in Appendix IV shows that women are less likely to search for jobs through employment agencies, contacting employers directly, or through internships.

Table 4: Likelihood of an Occupation-Education mismatch in STEP.

	...mismatch in current job		...mismatch in first job	
Marginal effect of each variable on likelihood of...	m1	m2	m3	m4
Female	-0.071*	-0.073*	-0.071*	-0.072*
	(0.034)	(0.034)	(0.035)	(0.036)
Age		-0.0010		-0.00055
		(0.0016)		(0.0017)
Province		*Yes*		*Yes*
Observations	850	850	764	764

* p<0.05, ** p<0.01, *** p<0.001

The coefficient on female dummy is negative in Table 4, indicating that women are less likely to be mismatched (i.e. more likely to be working in their field of study) either in their first job or in their current job. This result does not support the hypothesis that sorting across occupations occurs during transition to work due to involuntary occupation-education mismatches that are more likely for women.[14] However, we also acknowledge that our approach to testing this hypothesis comes with a number of limitations.

First, this exercise is sensitive to the level of aggregation in the data. As the level of aggregation increases, it becomes more difficult to identify when people are working outside their field of study because the fields of study and occupations become broader. We use the lowest level of aggregation available, which is occupation at the 3-digit level and field of education as available in the STEP data. For our analysis, the information on occupation is a level of aggregation higher than in the LFS and the fields of study in STEP, are broad and span several specializations. This renders our dummy for mismatch a conservative estimate of the true extent of mismatch. Repeating this analysis with more detailed information on fields of study could lead to different conclusions.

Second, we cannot make the distinction between mismatches that are involuntary and those that are voluntary. While our hypothesis is about the former, voluntary mismatches may be desirable. For instance, they may reflect reallocations of labor to more productive and higher paying occupations. Finally, our indicator of mismatch comes from a classification of occupations to fields of study. Like any

[14] We also test whether women are more likely to be overqualified for their jobs in Table 7 in Appendix IV. We do not find significant differences.

such classification, this is bound to be imperfect. Altogether, the results may be reflective of these limitations.

6. Conclusions

We establish that the gender wage gap persists in Vietnam despite the fact that women have achieved similar levels of education as men or even surpassed men at higher levels of education. We also demonstrate that sorting into different occupations and industries explains a non-trivial portion of this gap.

Although occupational sorting is a phenomenon common to labor markets, relatively little is known empirically about what drives occupational sorting and when it emerges. Our analysis is a first step in this direction. We offer three explanations of the emergence of occupational sorting and test these empirically using data from Vietnam. We find little support for a hypothesis of occupational sorting being driven by social norms, which affect aspirations at a young age and subsequently influence human capital accumulation decisions. We also do not find evidence for a hypothesis of occupational sorting occurring due to the gender specific differences in the school-to-work transition.

We do find support for the hypothesis that women have stronger preferences than men for favorable non-monetary job characteristics such as having a formal contract, social insurance, health insurance, paid leave and shorter weekly hours. This is in line with a broad body of literature starting from Becker (1985) which points to the unequal distribution of the burden of household responsibilities affecting women's labor force participation and earnings. Interestingly, we find that stronger preferences for favorable non-monetary job characteristics are reflected in the aspirations of girls at the age of 12.

These results point towards two streams of interventions. First, interventions that enable women to better balance household and market roles may prove effective. These include making childcare services more easily available, allowing flexible and gender-neutral parental leave policies (Jenkner and Correa, 2017) or allowing flexible work arrangements such as telecommuting (Dettling, 2013). The design of such interventions requires further attention. For instance, policies that mandate the provision of social services to women by their employers can increase the costs of hiring women and be counterproductive - Prada, Rucci and Urzua (2015) find that the starting wages of women were lower after the introduction of mandated employer-provided childcare in Chile in the firms that were required to comply with this legislation. Similarly, simply expanding the number of spots available for childcare may not be sufficient to have an impact on female labor force participation if it does not tackle the demand side constraints to accessing childcare. Taking the example of the Government of Chile's large plans for expansion of childcare services, Mateo-Diaz and Rodriguez-Chamussy point out that 9% of spots remained unused in the largest national program, Junta Nacional de Jardines Infantiles (JUNJI). They highlight the importance of factors such as affordability, distance to home, and opening and closing hours of the centers.

Second, changing cultural norms about household work to enable a more equitable distribution of the burden of household responsibilities will likely have impacts. Promundo, for instance, has a number of

programs to shift social norms and gender attitudes, including educating men in order to increase their participation in caregiving activities.[15]

Finally, our results on the persistence of the gender wage gap and the wage gap in different groups are also relevant in the context of the currently ongoing revision of Vietnam's labor code, expected to be finalized in 2019. The proposed changes include improvement to the protection of female employees and an increase in the retirement age of female and male workers (and a decrease of the gap in retirement age between genders). Our results provide empirical support for these changes from the perspective of their influence on the gender wage gap.

[15] See Promundo's website for more: http://promundoglobal.org/programs/mencare/.

7. References

Arcidiacono, Peter. "Ability Sorting and the Returns to College Major." *Journal of Econometrics* 121, no. 1-2 (2004): 343-75.

Babcock, Linda, Maria P. Recalde, Lise Vesterlund, and Laurie Weingart. "Gender Differences in Accepting and Receiving Requests for Tasks with Low Promotability." *American Economic Review* 107, no. 3 (2017): 714-47.

Barbulescu, Roxana, and Matthew Bidwell. "Do Women Choose Different Jobs from Men? Mechanisms of Application Segregation in the Market for Managerial Workers." *Organization Science* 24, no. 3 (2013): 737-56.

Becker, Gary S. "Investment in Human Capital: A Theoretical Analysis" *Journal of Political Economy* 70, no. 5 (1962): 9-49.

Becker, Gary S: *Human Capital Revisited*. Chicago: University of Chicago, 1989.

Bharadwaj, Prashant, Giacomo de Giorgi, David Hansen, and Christopher Neilson, "The Gender Gap in Mathematics: Evidence from Low- and Middle Income Countries," October 2012. NBER Working Paper 18464

Booth, Alison L., and Patrick Nolen. "Gender Differences in Risk Behaviour: Does Nurture Matter?." *The Economic Journal* 122, no. 558 (2012).

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. "Gender, Competitiveness and Career Choices." *The Quarterly Journal of Economics* 129, no. 3 (2014): 1409-447.

Cárdenas, Juan-Camilo, Anna Dreber, Emma Von Essen, and Eva Ranehill. "Gender Differences in Competitiveness and Risk Taking: Comparing Children in Colombia and Sweden." *Journal of Economic Behavior and Organization* 83, no. 1 (2010): 11-23.

Daymont, Thomas N., and Paul J. Andrisani. "Job Preferences, College Major, and the Gender Gap in Earnings." *The Journal of Human Resources* 19, no. 3 (1984): 408-28.

Dercon, Stefan, and Abhijeet Singh. "From Nutrition to Aspirations and Self-Efficacy: Gender Bias over Time among Children in Four Countries." *World Development* 45 (2013): 31-50.

Dreber, Anna, Emma Von Essen, and Eva Ranehill. "Outrunning the Gender Gap - Boys and Girls Compete Equally." *Experiential Economics* 14, no. 4 (2011): 567-82.

Fafchamps, Marcel, Måns Söderbom, and Najy Benhassine. "Wage Gaps and Job Sorting in African Manufacturing." *Journal of African Economies* 18, no. 5 (2009): 824-68.

Fryer, Roland, and Steven Levitt. "An Empirical Analysis of the Gender Gap in Mathematics." *American Economic Journal: Applied Economic* 2, no. 2 (2009): 210-40.

Ganguli, Ina, Ricardo Hausmann, and Martina Viarengo. *Closing the Gender Gap in Education: Does It Foretell the Closing of the Employment, Marriage, and Motherhood Gaps?* Working paper. 2011.

Gneezy, Uri, and Aldo Rustichini. "Gender and Competition at a Young Age." *American Economic Review* 94, no. 2 (2004): 377-81.

Goldin, Claudia. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104, no. 4 (2014): 1091-119.

Grant, Monica J., and Jere R. Behrman. "Gender Gaps in Educational Attainment in Less Developed Countries." *Population and Development Review* 36, no. 1 (2010): 71-89.

Groshen, Erica L. "The Structure of the Female/Male Wage Differential : Is It Who You Are, What You Do, or Where You Work?" *Journal of Human Resources* (1991): 457-72.

Hausmann, Ricardo, Laura D. Tyson, Yasmina Bekhouche, and Saadia Zahidi. "The Global Gender Gap 2012." In *World Economic Forum*. 2013.

Migheli, Matteo. "Gender at Work: Incentives and Self-sorting." *Journal of Behavioral and Experimental Economics* 55 (2015): 10-18.

Mincer, Jacob: *Schooling, Experience, and Earnings*, Columbia University Press: New York, (2004).

Montt, Guillermo. "The Causes and Consequences of Field-of-study Mismatch." (2015).

Nordin, Martin, Inga Persson, and Dan-Olof Rooth. "Education–occupation Mismatch: Is There an Income Penalty?" *Economics of Education Review* 29, no. 6 (2010): 1047-059.

Petersen, Trond, and Laurie A. Morgan. "Separate and Unequal: Occupation-Establishment Sex Segregation and the Gender Wage Gap." *American Journal of Sociology* 101, no. 2 (1995): 329-65.

Robst, John. "Education and Job Match: The Relatedness of College Major and Work." *Economics of Education Review* 26, no. 4 (2007a): 397-407.

Robst, John. "Education, College Major, and Job Match: Gender Differences in Reasons for Mismatch." *Education Economics* 15, no. 2 (2007b): 159-75.

World Bank. 2008. *Gender in Agriculture Sourcebook (English)*. Agriculture and Rural Development. Washington, DC: World Bank.

World Bank. 2012. *World Development Report 2012 : Gender Equality and Development.* Washington, DC: World Bank.

World Bank. 2012. *Toward Gender Equality in East Asia and the Pacific : A Companion to the World Development Report.* World Bank East Asia and Pacific Regional Report. Washington, DC: World Bank.

Zafar, Basit. "College Major Choice and the Gender Gap." *Journal of Human Resources* 43, no. 3 (2013): 545-95.

Zveglich, Joseph E., and Yana Van Der Meulen Rodgers. "Occupational Segregation and the Gender Wage Gap in a Dynamic East Asian Economy." *Southern Economic Journal* 70, no. 4 (2004): 850-75.

Appendix I: Gender gaps in educational enrollment and labor force participation in Vietnam

Figure 13: Ratio of female to male enrollment in primary education (2015).

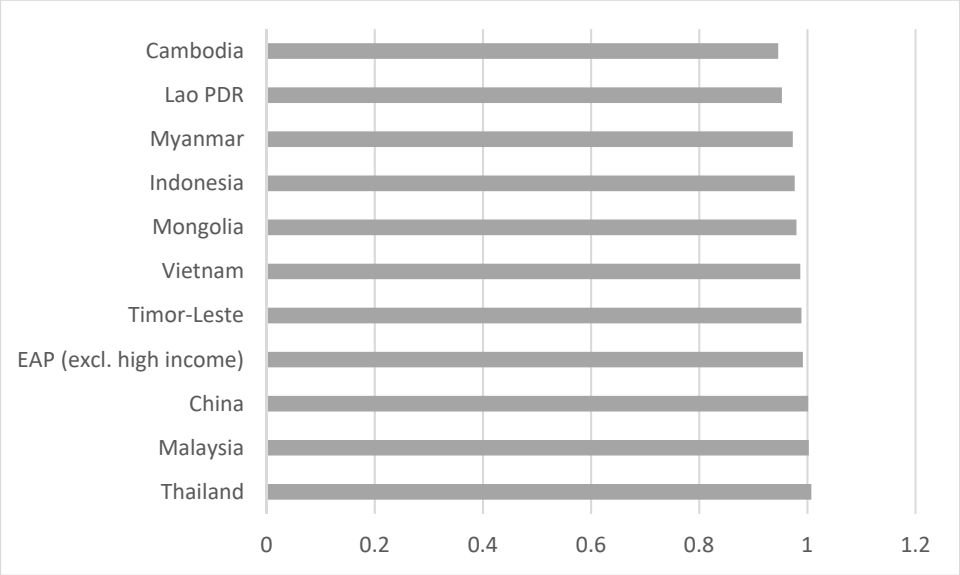
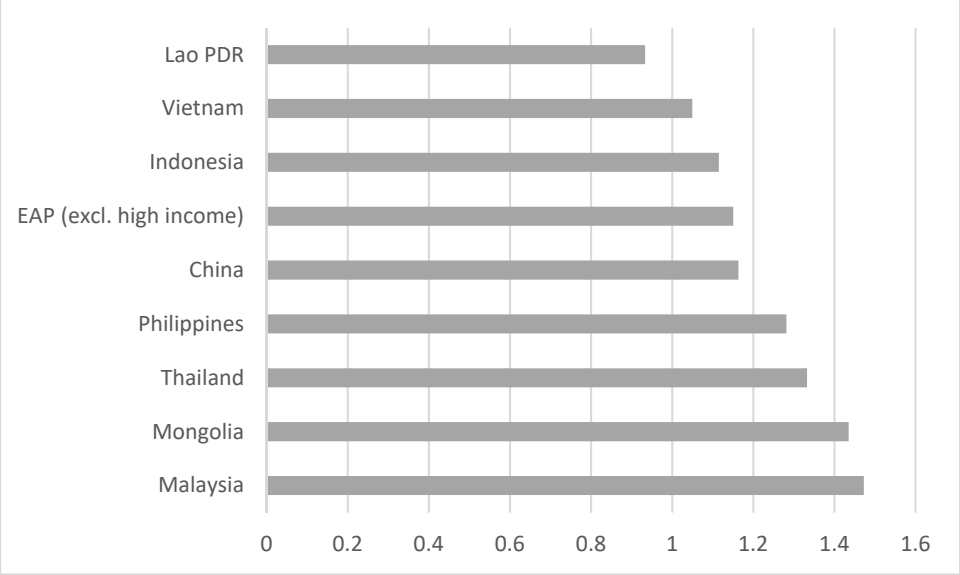
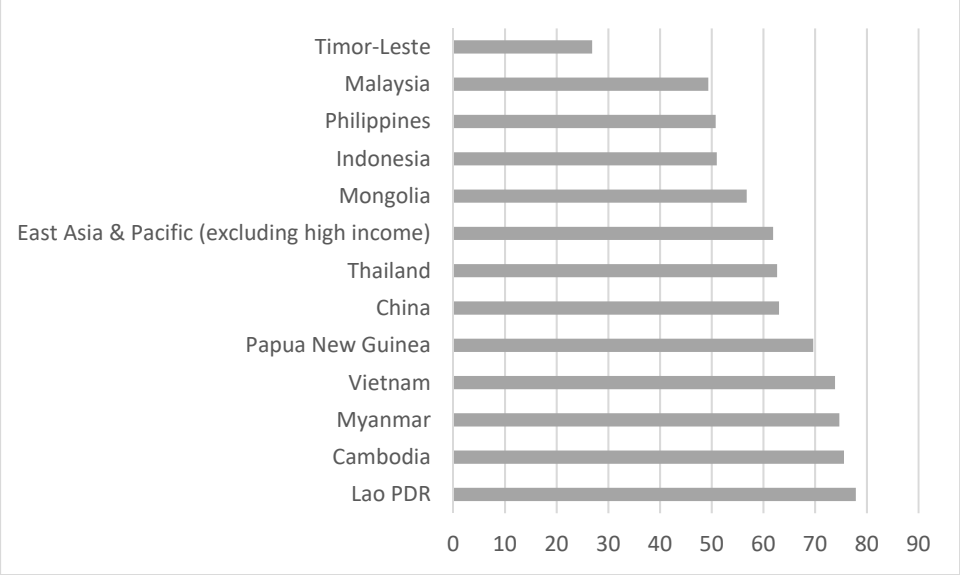


Figure 14: Ratio of female to male enrollment in tertiary education (2015).



Source (Fig 13, 14): United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics.

Figure 15: Female labor force participation rates (2017).



Source: International Labour Organization, ILOSTAT database. Data retrieved in March 2017.

Appendix II: LFS - description of the study sample

Table 5:Demographics (LFS))

	2011		2012		2013		2014	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Rural/Urban								
Urban	69.2%	73.7%	59.3%	64.3%	58.9%	64.0%	59.5%	63.9%
Rural	30.8%	26.3%	40.7%	35.7%	41.1%	36.0%	40.5%	36.1%
Married								
Not married	28.8%	30.3%	27.9%	29.0%	27.1%	27.7%	26.8%	26.5%
Married	71.2%	69.7%	72.1%	71.0%	72.9%	72.3%	73.2%	73.5%
Age Cohort								
15-19	5.2%	4.8%	5.0%	4.8%	4.5%	3.9%	4.0%	3.5%
20-24	12.4%	14.5%	12.2%	13.8%	12.1%	13.7%	11.8%	13.2%
25-29	17.6%	20.8%	16.8%	19.8%	16.5%	19.5%	16.1%	19.0%
30-34	15.5%	16.7%	15.3%	16.8%	15.8%	17.7%	16.3%	18.0%
35-39	13.6%	13.1%	14.1%	14.0%	13.9%	13.8%	13.9%	14.4%
40-44	11.7%	10.8%	11.8%	11.4%	11.9%	11.5%	12.1%	11.6%
45-49	10.3%	9.5%	10.5%	9.4%	10.0%	9.1%	10.2%	9.4%
50-54	8.1%	7.0%	8.0%	7.0%	8.4%	7.4%	8.3%	7.6%
55-59	3.9%	1.7%	4.4%	1.9%	4.8%	2.2%	5.2%	2.1%
60 and above	1.6%	1.0%	1.7%	1.2%	2.1%	1.3%	2.1%	1.3%
Highest Education								
ICSED 1: Primary education or lower	24.6%	21.2%	25.0%	21.7%	24.6%	21.4%	23.6%	20.1%
ICSED 2: Lower Secondary	20.5%	17.2%	20.7%	17.6%	20.6%	16.6%	20.0%	17.0%
Short vocational training course	5.4%	1.9%	6.3%	2.4%	7.3%	2.5%	6.6%	2.1%
ICSED 3: Upper secondary	26.0%	28.1%	24.9%	26.2%	23.4%	25.0%	23.7%	24.4%
ICSED 4: Post secondary non-tertiary	3.8%	9.2%	4.0%	9.3%	4.2%	9.7%	4.4%	9.4%
ICSED 5: Tertiary or higher	19.6%	22.4%	19.1%	22.8%	20.0%	24.8%	21.6%	26.9%

Table 6: Labor force status (LFS)

	2011		2012		2013		2014	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Labor force								
Out of labor force	18.7%	28.0%	18.0%	26.3%	16.8%	25.6%	16.6%	25.5%
In labor force	81.3%	72.0%	82.0%	73.7%	83.2%	74.4%	83.4%	74.5%
Employed								
Unemployed	3.0%	3.5%	2.5%	2.9%	3.0%	3.0%	2.9%	2.8%
Employed	97.0%	96.5%	97.5%	97.1%	97.0%	97.0%	97.1%	97.2%
Worker type (conditional on being employed)								
Employer	4.2%	2.2%	3.6%	1.7%	3.3%	1.5%	2.9%	1.4%
Wage worker	42.2%	32.6%	38.7%	29.4%	38.3%	29.6%	38.5%	30.2%
Unpaid	12.9%	23.5%	14.2%	25.3%	14.6%	25.0%	17.2%	27.3%
Self-employed	40.6%	41.7%	43.5%	43.5%	43.9%	43.9%	41.4%	41.1%
Type of enterprise (conditional on being a wageworker)								
Household or collective	38.4%	23.9%	39.20%	24.0%	40.0%	24.1%	39.0%	22.9%
Private sector	25.7%	30.5%	25.1%	30.5%	24.6%	31.1%	25.5%	33.0%
State or SOE	35.9%	45.6%	35.7%	45.5%	35.3%	44.8%	35.5%	44.1%

Appendix III – Simplified system of education in Vietnam

LFS

Classification of education levels in LFS	Simplified system used in the analysis
NEVER ATTENDED SOME PRIMARY PRIMARY EDUCATION	ICSED 1: Primary Education or less
LOWER SECONDARY	ICSED 2: Lower Secondary
SHORT -TERM TRAINING	Short-term vocational traning course
HIGHER SECONDARY TRADE VOCATIONAL SCHOOL VOCATIONAL SCHOOL	ICSED 3: Upper Secondary
TRADE COLLEGE COLLEGE	ICSED 4: Post-secondary Non-tertiary
UNIVERSITY AND OVER	ICSED 5: Tertiary or higher

STEP

Classification of education levels in LFS	Simplified system used in the analysis
PRIMARY SCHOOL	ICSED 1: Primary Education
LOWER SECONDARY SCHOOL	ICSED 2: Lower Secondary
ELEMENTARY VOCATIONAL	Short-term vocational training course
UPPER SECONDARY SCHOOL UPPER SECONDARY VOCATIONAL PROFESSIONAL SCHOOL	ICSED 3: Upper Secondary
VOCATIONAL COLLEGE COLLEGE (NON-VOCATIONAL)	ICSED 4: Post-secondary Non-tertiary
BACHELOR MASTER'S DOCTORATE/ PHD	ICSED 5: Tertiary or higher

Appendix IV: How women find their jobs (STEP)

Figure 16: Likelihood of women reporting each method, relative to men (marginal effects from Probit regression).

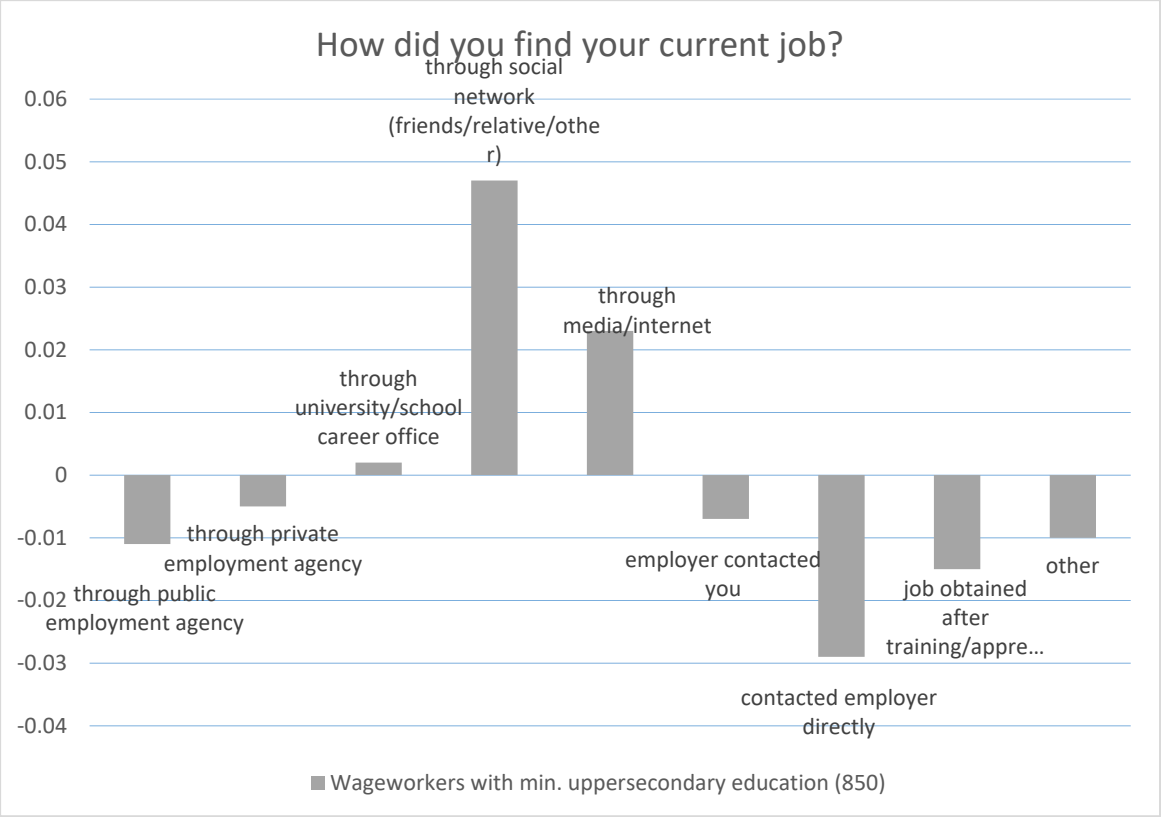


Table 7: Likelihood of being overqualified for current job (STEP)

	Probability of being overqualified		
	(1)	(2)	(3)
female	-0.018	-0.016	-0.027
	(0.03)	(0.03)	(0.04)
Age		-0.011	-0.011
		(0.01)	(0.01)
Age squared		0.000	0.000
		(0.00)	(0.00)
Field of Education			*Yes*
Observations	859	859	859